

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representation of
The original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification ⁵ : C08J 7/18, C08K 7/02</p>	<p>A1</p>	<p>(11) International Publication Number: WO 91/18047 (43) International Publication Date: 28 November 1991 (28.11.91)</p>
<p>(21) International Application Number: PCT/US91/03540 (22) International Filing Date: 20 May 1991 (20.05.91) (30) Priority data: 528,300 24 May 1990 (24.05.90) US (71) Applicant: GENENTECH, INC. [US/US]; 460 Point San Bruno Boulevard, South San Francisco, CA 94080 (US). (72) Inventors: HAMMONDS, R., Glenn ; 7036 Norfolk Road, Berkeley, CA 94705 (US). MASON, Anthony, J. ; 1446 Floribunda, #204, Burlingame, CA 94010 (US). (74) Agents: HASAK, Janet, E. et al.; Genentech, Inc., Legal Department, 460 Point San Bruno Boulevard, South San Francisco, CA 94080 (US).</p>		<p>(81) Designated States: AT (European patent), BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FR (European patent), GB (European patent), GR (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent). Published <i>With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i></p>
<p>(54) Title: MAMMALIAN EXPRESSION OF THE BMP-2 FAMILY</p>		
<p>(57) Abstract A DNA construct is provided comprising DNA encoding a mature BMP-2 upstream of which is DNA encoding a precursor portion of a mammalian protein other than the BMP-2. Also provided are mammalian expression vectors and hosts containing such a DNA construct and methods for improved expression using such construct.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MG	Madagascar
AU	Australia	FI	Finland	ML	Mali
BB	Barbados	FR	France	MN	Mongolia
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GN	Guinea	NL	Netherlands
BJ	Benin	GR	Greece	NO	Norway
BR	Brazil	HU	Hungary	PL	Poland
CA	Canada	IT	Italy	RO	Romania
CF	Central African Republic	JP	Japan	SD	Sudan
CG	Congo	KP	Democratic People's Republic of Korea	SE	Sweden
CH	Switzerland	KR	Republic of Korea	SN	Senegal
CI	Côte d'Ivoire	LI	Liechtenstein	SU	Soviet Union
CM	Cameroon	LK	Sri Lanka	TD	Chad
CS	Czechoslovakia	LU	Luxembourg	TG	Togo
DE	Germany	MC	Monaco	US	United States of America
DK	Denmark				

MAMMALIAN EXPRESSION OF THE BMP-2 FAMILY

Field of the Invention

This invention relates to an improved method for expressing DNA encoding the bone morphogenetic protein-2 family in mammalian cells.

5 Description of Related Art

The disorders associated with bone loss present major public health problems for Western societies. Osteoporosis alone may affect 20 million Americans in the early years of the next century. Hence, there is wide interest in identifying factors or potential therapeutic agents that inhibit bone loss and stimulate the formation of healthy new bone.

10 Bone is an extremely complex, but highly organized, connective tissue that is continuously remodeled during the life of an adult by cellular events that initially break it down (osteoclastic resorption) and then rebuild it (osteoblastic formation). This remodeling process occurs in discrete packets throughout the skeleton, i.e., in both cortical bone and trabecular bone. It has recently been reported that mouse bone marrow cells can be
15 stimulated to generate osteoclasts in the presence of parathyroid hormone-related protein or vitamin D. See Akatsu et al., Endocrinology, 125: 20-27 (1989); Takahashi et al., Endocrinology, 123: 2600-2602 (1988) and Takahashi et al., Endocrinology, 123: 1504-1510 (1988).

The currently available therapeutic agents known to stimulate bone formation are
20 fluoride, estrogen, metabolites, and vitamin D. Fluoride clearly increases trabecular bone mass, but questions remain about the quality of the new bone formed, the side effects observed in some patients, whether there are beneficial effects on vertebral fracture rates, and whether increased fragility of cortical bone with subsequent propensity to hip fracture follows.

25 Another approach is using agents that promote resorption (parathyroid hormone) and then interrupt resorption (calcitonin). One proposed, but not validated, such sequential therapeutic regimen is coherence therapy, where bone metabolic units are activated by oral phosphate administration and then resorption is inhibited by either diphosphonates or calcitonin.

30 Within the past few years several factors that stimulate osteoblasts have been identified in bone, including transforming growth factor- β (TGF- β), fibroblast growth factor, platelet-derived growth factor, insulin-like growth factor-I, and β 2 macroglobulin.

Other proteins stored in the bone matrix may also be important for bone formation. When demineralized bone was injected into the muscle or subcutaneous tissue of rats, a
35 cascade of events, including chondrogenesis, ensued. Urist, Science, 150: 893 (1965). Since the 1960s several investigators have attempted to identify and characterize this activity and have provided an assay for purification of such activity. Reddi and Huggins, Proc. Natl.

Acad. Sci. USA, 69: 1601-1605 (1972); Sampath and Reddi, Proc. Natl. Acad. Sci. USA, 78: 7599-7603 (1981).

This assay served as the basis for purifying several novel proteins from bone in sufficient quantity and purity to provide amino acid sequence information, including
5 osteogenin, a protein of 22 Kd [Sampath et al., Proc. Natl. Acad. Sci. USA, 84: 7109 (1987); Luyten et al., J. Biol. Chem., 264: 13377-13380 (1989)] and a glycoprotein called osteoinductive factor [Bentz et al., J. Cell. Biol., 107: 162a (1989)]. See also Wang et al., Proc. Natl. Acad. Sci., 85: 9484-9488 (1988). Based on amino acid sequence data, clones
10 encoding several proteins related by sequence similarity to TGF- β were isolated from bovine and human sources. Wozney et al., Science, 242: 1528-1534 (1988); PCT WO 88/00205 published January 14, 1988; U.S. 4,877,864 issued October 31, 1989. These latter proteins included BMP-2A (also known as BMP-2), BMP-2B (also known as BMP-4), and BMP-3. The sequence of tryptic peptides from osteogenin match the sequence reported for BMP-3.

The TGF- β supergene family includes five distinct forms of TGF- β [Sporn and Roberts,
15 in Peptide Growth Factors and Their Receptors, Sporn and Roberts, eds. (Springer-Verlag: Berlin, 1990) pp. 419-472], as well as the differentiation factors vg1-[Weeks and Melton, Cell, 51: 861-867 (1987)] and DPP-C polypeptide [Padgett et al., Nature, 325: 81-84 (1987)], the hormones activin and inhibin [Mason et al., Nature, 318: 659-663 (1985); Mason et al., Growth Factors, 1: 77-88 (1987)], the Mullerian-inhibiting substance, MIS [Cate et al., Cell, 45: 685-698 (1986)], the BMPs, and the developmentally regulated protein Vgr-1 [Lyons et al., Proc. Natl. Acad. Sci. USA, 86: 4554-4558 (1989)]. The subset BMP-2A and BMP-2B is approximately 75% homologous in sequence to DPP-C and may represent the mammalian equivalent of that protein.

The proteins of the TGF- β supergene family are disulfide-linked homo- or heterodimers
25 encoded by larger precursor polypeptide chains containing a hydrophobic signal sequence, a long and relatively poorly conserved N-terminal pro region of several hundred amino acids, a cleavage site (usually polybasic), and a shorter and more highly conserved C-terminal region. This C-terminal region corresponds to the processed mature protein and contains approximately 100 amino acids with a characteristic cysteine motif, i.e., the conservation of
30 seven of the nine cysteine residues of TGF- β among all known family members. Although the position of the cleavage site between the mature and pro regions varies among the family members, the C-terminus of all of the proteins is in the identical position, ending in the sequence Cys-X-Cys-X, but differing in every case from the TGF- β consensus C-terminus of Cys-Lys-Cys-Ser. Sporn and Roberts, 1990, *supra*.

35 The pro region of TGF- β associates non-covalently with the mature TGF- β dimer [Wakefield et al., J. Biol. Chem., 263: 7646-7654 (1988); Wakefield et al., Growth Factors, 1: 203-218 (1989)], and the pro regions are found to be necessary for proper folding and secretion of the active mature dimers of both TGF- β and activin [Gray and Mason, Science,

247: 1328-1330 (1990)]. The association between the mature and pro regions of TGF- β masks the biological activity of the mature dimer, resulting in formation of an inactive latent form. Latency is not a constant of the TGF- β supergene family, since the presence of the pro region has no effect on activin or inhibin biological activity.

5 A unifying feature of the biology of the proteins from the TGF- β supergene family is their ability to regulate developmental processes. Regarding bone formation *in vivo*, of all the proteins in the TGF- β supergene family, the BMPs and TGF- β play the most major role.

Recombinant TGF- β 1 has been cloned [Derynck et al., Nature, 316:701-705 (1985)] and expressed in Chinese hamster ovary cells [Gentry et al., Mol. Cell. Biol., 7: 3418-3427 (1987)]. Additionally, recombinant human TGF- β 2 [deMartin et al., EMBO J., 6: 3673 (1987)], as well as human and porcine TGF- β 3 [Derynck et al., EMBO J., 7: 3737-3743 (1988); ten Dijke et al., Proc. Natl. Acad. Sci. USA, 85: 4715 (1988)], have been cloned. Expression levels of the mature TGF- β 1 protein in COS cells are increased by substituting a serine residue for cysteine residues located in the pro region of the TGF- β 1 precursor.

10
15 Brunner et al., J. Biol. Chem., 264: 13660-13664 (1989)].

BMP-2A and BMP-3 have been recombinantly produced in monkey COS-1 cells and Chinese hamster ovary cells by Wozney et al., *supra*. However, the level of expression of BMP-2A and -2B cDNA is relatively low when the DNA is not amplified. Higher levels of BMP-2A protein expression in CHO cells have been obtained by amplification to a high copy number using methotrexate selection of dihydrofolate reductase. Wang et al., Proc. Natl. Acad. Sci. USA, 87: 2220-2224 (1990).

20

Confirmation of the osteogenic activity of BMPs and commercial production thereof depend on the ability to produce useful amounts of active material by recombinant means of expression and development of methods to purify them in an active form. The ability to successfully reconstitute endochondral bone formation remains the standard by which to judge the osteogenic character of candidate factors. The biological activities of BMP-2A, BMP-3, and an unrelated molecule, BMP-1, were originally assessed in an implant model using material expressed in COS cells, resulting in only cartilage formation. Wozney et al., *supra*. More recently, the partially purified BMP-2A expressed in CHO cells was shown to require a dose of at least 600 ng/implant to induce cartilage and bone formation. Wang et al., 1990, *supra*. The osteogenic activities of BMP-2B and BMP-3 have not been established.

25
30

It is an object of the present invention to provide purified BMP-2B in sufficient quantities to test for its osteogenic activity, and to produce it on a commercial scale.

It is another object to improve the expression levels of BMP-2 DNA in mammalian cells without amplifying the DNA.

35

It is still another object to achieve higher production of BMP-2 protein than was previously attained at a level of amplification equivalent to that previously employed.

These and other objects will be apparent to those of ordinary skill in the art of molecular biology.

Summary of the Invention

Accordingly, this invention provides a DNA construct comprising DNA encoding a mature BMP-2 upstream of which is DNA encoding a precursor portion of a mammalian protein other than that of BMP-2. Preferably, the precursor portion has at least 25% amino acid sequence identity to the native precursor portion of the BMP-2 in the region spanning the N-terminus of the BMP-2 precursor to the first cysteine residue in the mature BMP-2.

In another aspect, this invention provides an expression vector comprising the above-described DNA construct and hosts transformed with such a vector.

In a method for expressing DNA encoding a BMP-2 in mammalian cells, this invention also furnishes the improvement which comprises employing as the host the host transformed with the vector described above.

Additionally, this invention provides a method for producing BMP-2 by culturing mammalian host cells transfected with the expression vector described above, the cells being capable of expressing the DNA construct of the vector, and recovering mature BMP-2 from the cells. Preferably, the recovery is from the host cell media (in which case the expression vector contains a signal sequence, whether native to the precursor or BMP-2 or heterologous to the precursor or BMP-2, that directs secretion of the mature BMP-2 to the medium).

The result of this method is dramatically improved expression levels of BMP-2 DNA in mammalian cells over that attainable using the BMP-2 precursor portion that is native to the BMP-2 to be produced.

Brief Description of the Drawings

Figure 1 depicts the amino acid sequences of BMP-2A and BMP-2B and indicates the regions of sequence identity. The junction between the precursor portions and mature portions is shown by a vertical line with two arms.

Figure 2 depicts the complete amino acid sequence of the chimera of the precursor portion of BMP-2A and the mature region of BMP-2B.

Figure 3A depicts expression plasmid pRK5.bmp2/4-1.1, and Figure 3B depicts the junction region of the BMP-2A/2B hybrid insert. A portion of an alignment of BMP-2A and BMP-2B is shown with identical residues boxed. The coding sequence resulting from fusion of BMP-2A and BMP-2B is shaded showing the crossover point. The underlined sequence with an arrow indicates sequence confirmed by Edman degradation of purified recombinant BMP-2B.

Figure 4 depicts a fluorogram of an SDS-PAGE reducing gel of supernatants from human embryonic kidney cell line transfections with DNA encoding either the native BMP-2A molecule (lane 1), the chimeric BMP-2A/2B molecule (lane 2), the native BMP-2B molecule (lane 3), control pRK5 plasmid (lane 4), or no plasmid (lane 5). Figure 5 depicts graphs

of calcium content (Fig. 5A) and alkaline phosphatase content (Fig. 5B) of implants in rats (harvested at 12 days) of demineralized bone powder (DBP) or guanidine-HCl-extracted DBP reconstituted with the indicated amounts of mature recombinant BMP-2B or TGF- β . The solid and cross-hatched bars presented for two doses are duplicate runs.

5

Description of the Preferred Embodiments

Definitions

As used herein, the term "BMP-2" refers to the family of bone morphogenetic proteins of the type 2, derived from any species. Reference to BMP-2 herein is understood to be a reference to any one of the currently identified forms, including BMP-2A and BMP-2B (formerly called BMP-4) described by Wozney et al., *supra*, and WO 88/00205, *supra*, the sequences of which are shown in Figure 1, as well as to BMP-2 species identified in the future. The term "BMP-2" also includes polypeptides derived from the sequence of any known BMP-2 whose mature sequence is at least about 75% homologous with the sequence of a mature BMP-2, including DPP-C. Members of the BMP-2 family appear to be encoded as a larger precursor that shares a region of high homology near the N-terminus.

15

As used herein, "precursor portion" refers to the polypeptide sequence derived from a prepro-mammalian protein representing either the pro-domain or prepro-domain without the mature protein. Candidate mammalian proteins having such precursor portions are those encoded as larger precursors that typically contain a signal sequence at their N-terminus followed by a dibasic amino acid cleavage site and a pro-region, followed by another dibasic amino acid cleavage site and the mature region of the protein. Thus, the precursor portion is that which is N-terminal to the mature N-terminus of the mammalian protein and may include the signal sequence for secretion of that protein. Preferably, the mammalian protein from which the precursor portion is derived is a member of the TGF- β supergene family, as described above. Examples of suitable precursor portions are those wherein the signal sequence is followed by a sequence that represents a polypeptide region that after cleavage reassociates with the mature protein covalently or non-covalently, as in the case of insulin, relaxin, inhibin, activin, and TGF- β .

20

25

The expression "at least 25% amino acid sequence identity to the native precursor portion of the BMP-2 from the N-terminus of the BMP-2 precursor to the first cysteine residue in the mature region of the BMP-2" refers to a precursor portion that shares this minimum sequence identity to the relevant portion of the BMP-2 DNA being expressed. This sequence identity can be readily calculated for BMP-2A and BMP-2B from the entire amino acid sequences shown in Figure 1. As examples, the precursor portion of BMP-2A shares 55% amino acid sequence identity to the native precursor portion of BMP-2B from the N-terminus of the BMP-2B precursor to the first cysteine residue in the mature region of the BMP-2B molecule, and vice-versa. The precursor of the protein vgr [Lyons et al., *Proc. Natl. Acad. Sci. USA*, 86: 4554-4558 (1989)], which is related to the product of an amphibian gene vg1

30

35

expressed in frog oocytes, shares 25% homology with the relevant portion of BMP-2B. The protein decapenta-plegic gene complex from *Drosophila*, DPP-C [Padgett et al., *Nature*, 325: 81-84 (1987)], shares 27% and 28% amino acid sequence identity with the relevant portions of BMP-2A and BMP-2B, respectively. Most preferred herein is the use of the BMP-2A prepro-domain as the precursor portion for secreting mature BMP-2B.

Modes for Carrying Out the Invention

The vectors and methods disclosed herein are suitable for use for expression in a wide range of mammalian host cell lines.

In general, prokaryotes such as, e.g., *E. coli* strains are preferred for cloning, amplifying, or storing the vectors of interest. Vector DNA is easily obtainable from certain prokaryotes. *E. coli* K12 strain MM 294 (ATCC No. 31,446) is particularly useful for this purpose, as are *E. coli* B and *E. coli* X1776 (ATCC No. 31,537). In general, plasmid vectors containing replicon and control sequences that are derived from species compatible with the host cell are used in connection with these prokaryotic hosts. The vector ordinarily carries a replication site, as well as marking sequences that are capable of providing phenotypic selection in transformed cells. For example, *E. coli* is typically transformed using pBR322, a plasmid derived from an *E. coli* species [see, e.g., Bolivar et al., *Gene*, 2: 95 (1977)]. pBR322 contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR322 plasmid, or other microbial plasmid or phage, must also contain, or be modified to contain, promoters that can be used by the microbial organism for expression of the selectable marker genes.

Cultures of cells derived from mammalian organisms are useful as expression hosts using tissue culture methods [Tissue Culture, Academic Press, Kruse and Patterson, editors (1973)]. Examples of such useful host cell lines include monkey kidney CVI line transformed by SV40 sequences (COS-7, ATCC CRL 1651); human embryonic kidney line [293, Graham et al., *J. Gen. Virol.*, 36: 59 (1977)]; baby hamster kidney cells (BHK, ATCC CCL 10); Chinese hamster ovary cells [CHO, Urlaub and Chasin, *Proc. Natl. Acad. Sci. USA*, 77: 4216 (1980)]; mouse sertoli cells [TM4, Mather, *Biol. Reprod.*, 23: 243-251 (1980)]; monkey kidney cells (CVI, ATCC CCL 70); African green monkey kidney cells (VERO-76, ATCC CRL-1587); human cervical carcinoma cells (HELA, ATCC CCL 2); canine kidney cells (MDCK, ATCC CCL 34); buffalo rat liver cells (BRL 3A, ATCC CRL 1442); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); mouse mammary tumor cells (MMT 060562, ATCC CCL51); rat hepatoma cells [HTC, M1.54, Baumann et al., *J. Cell. Biol.*, 85: 1-8 (1980)]; and TRI cells [Mather et al., *Annals N.Y. Acad. Sci.*, 383: 44-68 (1982)]. The most preferred mammalian hosts herein are CHO and 293 cell lines.

Expression vectors for such cells ordinarily will contain control regions, which are specific sequences at the 5' and 3' ends of eukaryotic genes that may be involved in the control of either transcription, RNA processing, or translation. At the 3' end of most

eukaryotic genes is an AATAAA sequence that signals processing of the mRNA for polyadenylation addition.

Thus, the vector will typically include a promoter located in front of the gene to be expressed, polyadenylation sites, and transcriptional terminator sequences, all described in further detail herein. The vector may optionally also include an origin of replication. Further, the vector may contain, after the promoter, a transcription initiation site located in front of an optional splice unit, which is in turn located before the encoding gene.

Examples of suitable mammalian expression vectors are found in EP 307,247; 260,148; 309,237; and 307,248.

For use in mammalian cells, the control functions on the expression vectors are often provided by viral material. For example, commonly used promoters are derived from the genomes of polyoma, Adenovirus2, retroviruses, cytomegalovirus, and Simian Virus 40 (SV40). Other promoters are those from heterologous sources, e.g., the beta actin promoter. The early and late promoters of SV40 virus are particularly useful because both are obtained easily from the virus as a fragment that also contains the SV40 viral origin of replication [Fiers et al., Nature, 273: 113 (1978)]. Smaller or larger SV40 fragments may also be used, provided there is included the approximately 250-bp sequence extending from the HindIII site toward the BglI site located in the viral origin of replication. The immediate early promoter of the human cytomegalovirus is conveniently obtained as a HindIII restriction fragment. Greenaway et al., Gene, 18: 355-360 (1982). Further, it is also possible, and often desirable, to utilize promoter or control sequences normally associated with the desired gene sequence, provided such control sequences are compatible with the host cell systems.

Transcription of a DNA encoding a desired heterologous polypeptide by higher eukaryotes is increased by inserting an enhancer sequence into the vector. The enhancer is a cis-acting element of DNA, usually about from 10 to 300 bp, that acts on a promoter to enhance its transcription-initiation activity. Enhancers are relatively orientation and position independent, having been found 5' [Laimins et al., Proc. Natl. Acad. Sci. USA, 78: 993 (1981)] and 3' [Lusky et al., Mol. Cell Bio., 3: 1108 (1983)] to the transcription unit, within an intron [Banerji et al., Cell, 33: 729 (1983)] as well as within the coding sequence itself [Osborne et al., Mol. Cell Bio., 4: 1293 (1984)]. Preferably, however, the enhancer element is located upstream of the promoter sequence for this invention. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α -fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. One preferred enhancer is the SV40 enhancer region.

Expression vectors used in mammalian host cells will also contain polyadenylation sites. Examples of polyadenylation regions are those derived from viruses such as, e.g., the SV40 (early and late) or HBV.

5 An origin of replication may be provided either by construction of the vector to include an exogenous origin, such as may be derived from SV40 or other viral (e.g., Polyoma, Adeno, VSV, BPV) source, or may be provided by the host cell. If the vector is integrated into the host cell chromosome, the latter is often sufficient.

10 The expression vectors may suitably contain a selection gene, also termed a selectable marker. A selection gene encodes a protein necessary for the survival or growth of a host cell transformed with the vector. Examples of suitable selectable markers for mammalian cells include dihydrofolate reductase (DHFR), thymidine kinase (TK), or neomycin. When such selectable markers are successfully transferred into a mammalian host cell, the transformed mammalian host cell can survive if placed under selective pressure.

15 There are two widely used distinct categories of selective regimes. The first category is based on the metabolism of a cell and the use of a mutant cell line that lacks the ability to grow independent of a supplemented medium. Two examples are CHO.DHFR cells and mouse LTK cells. These cells lack the ability to grow without the addition of such nutrients as thymidine or hypoxanthine. Because these cells lack certain genes necessary for a complete nucleotide synthesis pathway, they cannot survive unless the missing nucleotides are provided in a supplemented medium. An alternative to supplementing the medium is to introduce an intact DHFR or TK gene into cells lacking the respective genes, thus altering their growth requirements. Individual cells that were not transformed with the DHFR or TK gene will not be capable of survival in non-supplemented medium. Therefore, direct selection of those cells requires cell growth in the absence of supplemental nutrients.

25 The second category is dominant selection, which refers to a selection scheme that does not require the use of a mutant cell line. This method typically employs a drug to arrest growth of a host cell. Those cells that have a novel gene would express a protein conveying drug resistance and would survive the selection. Examples of drugs used in dominant selection include neomycin [Southern and Berg, *J. Molec. Appl. Genet.*, 1: 327 (1982)], mycophenolic acid [Mulligan and Berg, *Science*, 209: 1422 (1980)], or hygromycin [Sugden et al., *Mol. Cell. Biol.*, 5: 410-413 (1985)]. The three examples given above employ bacterial genes under eukaryotic control to convey resistance to the appropriate drug, i.e., neomycin (G418 or geneticin), xgpt (mycophenolic acid), or hygromycin, respectively.

35 Extremely good amounts of polypeptide are produced by transiently transfected cell cultures using the method of this invention. It is also expected that stable transformants would result in higher production levels of the BMP-2 than transformants with the native proBMP-2 sequence. Furthermore, the process herein is expected to enhance production levels further when the cells are cotransfected with a separate vector encoding a secondary

coding sequence. One secondary coding sequence comprises dihydrofolate reductase (DHFR) that is affected by an externally controlled parameter, such as methotrexate (MTX), thus permitting control of expression by control of the MTX concentration.

Typical Methodology Employable

5 Construction of suitable vectors containing the desired coding and control sequences employs standard recombinant techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and re-ligated to form the desired plasmid.

If flush ends are required, the cleaved DNA preparation may be treated for 30 minutes at 37°C with DNA Polymerase I (Klenow fragment) or T4 DNA polymerase, phenol-chloroform
10 extracted, and ethanol precipitated. 3' protruding ends are removed by the 3' to 5' exonucleolytic activity of either enzyme, and the 5' protruding ends are made flush by the 5' to 3' polymerase activity incorporating complementary nucleotide until the end of the fragment is reached.

Size separation of the cleaved fragments may be performed using 6 percent
15 polyacrylamide gel described by Goeddel et al., Nucleic Acids Res., 8: 4057 (1980).

For analysis to confirm correct sequences in plasmids constructed, the ligation mixtures are typically used to transform *E. coli* K12 strain 294 (ATCC 31,446) or other suitable *E. coli* strains, and successful transformants selected by ampicillin or tetracycline resistance where appropriate. Plasmids from the transformants are prepared and analyzed by restriction
20 mapping and/or DNA sequencing by the method of Messing et al., Nucleic Acids Res., 9: 309 (1981) or by the method of Maxam et al., Meth. Enzym., 65: 499 (1980).

If amplification of the sequences is desired, DHFR-protein-coding DNA sequences are introduced into the mammalian cell host and stable transfectants are selected in the medium. The host cell cultures are grown in the presence of approximately 200-500 nM
25 concentrations of methotrexate, a competitive inhibitor of DHFR activity. The effective range of concentration is highly dependent, of course, upon the nature of the DHFR gene and the characteristics of the host. Clearly, generally defined upper and lower limits cannot be ascertained. Suitable concentrations of other folic acid analogs or other compounds that inhibit DHFR could also be used. MTX itself is, however, convenient, readily available, and
30 effective.

In order to simplify the examples and claims, certain frequently occurring methods will be referenced by shorthand phrases.

"Transfection" refers to the taking up of an expression vector by a host cell whether or not any coding sequences are in fact expressed. Numerous methods of transfection are
35 known to the ordinarily skilled artisan, for example, CaPO₄ and electroporation. Successful transfection is generally recognized when any indication of the operation of this vector occurs within the host cell.

"Transformation" means introducing DNA into an organism so that the DNA is replicable, either as an extrachromosomal element or by chromosomal integrant. Depending on the host cell used, transformation is done using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described by Cohen, S.N. Proc. Natl. Acad. Sci. (USA), 69: 2110 (1972); Mandel et al., J. Mol. Biol. 53:154 (1970); and more recently Liljestrom et al., Gene, 40: 241-246 (1985), is generally used for prokaryotes or other cells that contain substantial cell-wall barriers. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52: 456-457 (1978) is preferred. General aspects of mammalian cell host system transformations have been described by Axel in U.S. Pat. No. 4,399,216 issued August 16, 1983. Transformations into yeast are typically carried out according to the method of Van Solingen, et al., J. Bact., 130: 946 (1977) and Hsiao, et al., Proc. Natl. Acad. Sci. (USA) 76: 3829 (1979). However, other methods for introducing DNA into cells such as by nuclear injection or by protoplast fusion may also be used.

"Operably linked" refers to juxtaposition such that the normal function of the components can be performed. Thus, a coding sequence "operably linked" to control sequences refers to a configuration wherein the coding sequence can be expressed under the control of these sequences and wherein the DNA sequences being linked are contiguous and, in the case of a secretory leader, contiguous and in reading phase. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it effects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, then synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

"Control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, a ribosome binding site, and possibly, other as yet poorly understood sequences. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

"Expression system" refers to DNA sequences containing a desired coding sequence and control sequences in operable linkage, so that hosts transformed with these sequences are capable of producing the encoded proteins. To effect transformation, the expression system may be included on a vector; however, the relevant DNA may then also be integrated into the host chromosome.

As used herein, "cell," "cell line," and "cell culture" are used interchangeably and all such designations include progeny. Thus, "transformants" or "transformed cells" includes

the initial transformant and cultures derived therefrom without regard for the number of transfers. It is also understood that all progeny may not be precisely identical in DNA content, due to deliberate or inadvertent mutations. Mutant progeny that have the same functionality as screened for in the originally transformed cell are included. Where distinct
5 designations are intended, it will be clear from the context.

"Plasmids" are designated by a lower case p preceded and/or followed by capital letters and/or numbers. The starting plasmids herein are commercially available, are publicly available on an unrestricted basis, or can be constructed from such available plasmids in accord with published procedures. In addition, other equivalent plasmids are known in the
10 art and will be apparent to the ordinary artisan.

"Digestion" of DNA refers to catalytic cleavage of the DNA with an enzyme that acts only at specific nucleotide sequences in the DNA. Such enzymes are called restriction enzymes, and the sequence for which each is specific is called a restriction site. The various restriction enzymes used herein are commercially available and their reaction conditions, cofactors, and other requirements as established by the enzyme suppliers are used.
15 Restriction enzymes commonly are designated by abbreviations composed of a capital letter followed by other letters representing the microorganism from which each restriction enzyme originally was obtained and then a number designating the particular enzyme. In general, about 1 μ g of plasmid or DNA fragment is used with about 1-2 units of enzyme in about 20
20 μ l of buffer solution. Appropriate buffers and substrate amounts for particular restriction enzymes are specified by the manufacturer. Incubation of about 1 hour at 37°C is ordinarily used, but may vary in accordance with the supplier's instructions. After incubation, protein is removed by extraction with phenol and chloroform, and the digested nucleic acid is recovered from the aqueous fraction by precipitation with ethanol. When appropriate,
25 digestion with a restriction enzyme is followed by bacterial alkaline phosphatase-mediated hydrolysis of the terminal 5' phosphates to prevent the two ends of a DNA fragment from "circularizing," or forming a closed loop that would impede insertion of another DNA fragment at the restriction site. Unless otherwise stated, digestion of plasmids is not followed by 5' terminal dephosphorylation. Procedures and reagents for dephosphorylation are conventional
30 [Maniatis et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory, 1982) pp. 133-134].

"Recovery" or "isolation" of a given fragment of DNA from a restriction digest means separation of the digest on polyacrylamide or agarose gel by electrophoresis, identification of the fragment of interest by comparison of its mobility versus that of marker DNA
35 fragments of known molecular weight, removal of the gel section containing the desired fragment, and separation of the gel from DNA. This procedure is known generally. For example, see R. Lawn et al., Nucleic Acids Res. 9: 6103-6114 (1981), and D. Goeddel et al., Nucleic Acids Res. 8: 4057 (1980).

"Ligation" refers to the process of forming phosphodiester bonds between two double-stranded nucleic acid fragments [T. Maniatis et al., 1982, *supra*, p. 146]. Unless otherwise provided, ligation may be accomplished using known buffers and conditions with 10 units of T4 DNA ligase ("ligase") per 0.5 μ g of approximately equimolar amounts of the DNA fragments to be ligated.

"Preparation" of DNA from transformants means isolating plasmid DNA from microbial culture. Unless otherwise provided, the alkaline/SDS method of Maniatis et al., 1982, *supra*, p. 90, may be used.

"Oligonucleotides" are short-length, single- or double- stranded polydeoxynucleotides that are chemically synthesized by known methods [such as phosphotriester, phosphite, or phosphoramidite chemistry, using solid-phase techniques such as described in EP Pat. Pub. No. 266,032 published May 4, 1988, or via deoxynucleoside H-phosphonate intermediates as described by Froehler et al., Nucl. Acids Res., 14: 5399-5407 (1986)]. They are then purified on polyacrylamide gels.

The following example is intended to illustrate specific embodiments now known for practicing the invention, but the invention is not to be considered limited thereto.

EXAMPLE 1

cDNAs for BMP-2A and BMP-2B were cloned from a human placental cDNA library constructed in lambda gt10 [Ulrich et al., Nature, 313: 756-761 (1985)] using oligonucleotide probes based on the human nucleotide sequence [Wozney et al., *supra*] using standard cloning techniques [Sambrook et al., Molecular Cloning: A Laboratory Manual, Second Ed. (Cold Spring Harbor Laboratory, New York, 1989)]. The probes employed were as follows (where the initiator ATG is underlined and the direction from left to right is 5' to 3'):

25 BMP-2A Probes

CGACCATGGTGGCCGGGACCCGCTGTCTTCTAGCGTTGCTGCTTCCCCAGGTCCTCCTGG
GCGGCGCG (for 5' end)

AATGAAAAGGTTGTATTAAGA AACTATCAGGACATGGTTGTGGAGGGTTGTGGGTGTCGC
(for 3' end)

30 BMP-2B Probes

ATGATTCTGGTAACCGAATGCTGATGGTCGTTTTATTATGCCAAGTCCTGCTAGGAGGC
GCGAGCCATGCTAGTTTG (for 5' end)

CAGGAGATGGTAGTAGAGGGATGTGGGTGCCGCTGAGATCAGGCAGTCCTTGAGGATAG
ACAG (for 3' end)

35 No clones for BMP-3 were found in the human placental cDNA library using a similar approach to that above. Several cell lines were screened for expression of BMP-3 RNA by polymerase chain reaction amplification of the RNA [Mullis et al., Cold Spring Harbor Symp. Quant. Biol., 51: 263-273 (1986)] using oligonucleotide primers based on the human

nucleotide sequence [Wozney et al., *supra*]. One positive cell line, the NCI-H69 human small-cell lung carcinoma [Gazdar et al., *Cancer Res.*, 40: 3502-3507 (1980)] was identified. A cDNA library was prepared from the mRNA and screened with oligonucleotide probes using standard techniques (Sambrook et al., *supra*). The probe sequences were as follows (where the direction from left to right is 5' to 3'):

5 AGTGTCCCGCAGCGACGCCGGGAGCCGACGCGCCGCGGGTACCTAGCC (for 5' end)
TACCCTAACATGACAGTAGAGTCTTGCGCTTGCAGATAACCTGGCAAAGA (for 3' end)

Positive lambda gt10 clones were identified for BMP-2A, BMP-2B, and BMP-3 proteins, and these clones were sequenced. The sequenced clones encoding the BMP-2A and BMP-2B full-length proteins were digested with Sall. The expression vector pRK5 [EP 307,247 published 3/15/89] was also digested with Sall and the gel-isolated large fragment was ligated with the cDNA Sall digests encoding each BMP to create the expression plasmids pRK5.bmp2a and pRK5.bmp2b, for BMP-2A and BMP-2B, respectively.

The sequenced clone encoding the BMP-3 full-length protein was digested with EcoRI. pRK5 was also digested with EcoRI and the gel-isolated large fragment was ligated with the cDNA EcoRI digest encoding BMP-3 to create the expression plasmid pRK5.bmp3.

A human embryonic kidney cell line (293) [Graham et al., *supra*] was grown to confluence on 60-mm plates in F12:DMEM (1:1) medium (Gibco) containing 10% fetal calf serum (FCS) and transfected with one of the three BMP expression plasmids by the calcium phosphate method [Gorman, *DNA Cloning*, Vol. II (ed. Glover, D.), 143-190 (IRL, Oxford, 1985)]. More specifically, 5-10 μ g of one of the three BMP plasmid DNAs was mixed with 1 μ g of DNA encoding the VA RNA gene (Thimmappaya et al., *Cell*, 31: 543 (1982)) and dissolved in 250 μ l of 0.25 M CaCl₂. Added to this (dropwise while vortexing) was 250 μ l of 50 mM HEPES (pH 7.35), 280 mM NaCl, 1.5 mM NaPO₄, and the precipitate was allowed to form for 5-10 min. at 25°C. The suspended precipitate was then added to the cells and allowed to settle for 4-5 hours in the incubator. The medium was then aspirated off, the cell layer was washed with 5 ml of F12:DMEM (1:1), and 0.5 ml of 20% glycerol in phosphate-buffered saline (PBS) was added for 30 sec. A total of 5 ml of F12:DMEM (1:1) containing 10% fetal bovine serum was added, aspirated off, and replenished. 24 to 48 hours later, the 10% fetal bovine serum medium was replaced with serum-free F12:DMEM (1:1) minus cysteine and methionine. The cells were incubated for 2 hours at 37°C in 5% CO₂ in the presence of 200 μ Ci/ml ³⁵S-cysteine and 200 μ Ci/ml ³⁵S-methionine. Then the cell layers were washed with PBS and F12:DMEM (1:1) containing cysteine and methionine was added and the cells were allowed to incubate for 5-7 hours. Conditioned medium was then collected, concentrated 5-fold by lyophilization, and loaded on a 15% SDS gel, which was soaked with Enhance[®] (New England Nuclear) gel scintillation fluid, dried, and exposed to film at -80°C for 12 hours. Metabolic labeling of the conditioned medium revealed detectable

levels of expression that were low as compared with transfections of similar vectors containing activin or TGF- β cDNAs.

5 Conditioned medium from the cells transfected with BMP-2A, BMP-2B, or BMP-3 was partially purified by heparin-Sepharose chromatography as follows. A 5-ml heparin-Sepharose CL6B (Pharmacia) column was initially equilibrated with 4 M urea, 20 mM TrisCl at pH 7. Then the conditioned medium in 4 M urea, 20 mM TrisCl, pH 7, was loaded on the column. After loading, the fractions were eluted stepwise with 0, 0.1, 0.5 and 2.0 M NaCl in 4 M urea, 20 mM TrisCl, pH 7. The bone-forming activity of the fractions of each step was assessed *in vivo* by the method of Sampath and Reddi, *supra*. Both BMP-2A and BMP-2B
10 possessed easily demonstrable activity, but BMP-3 activity was more difficult to demonstrate. Not all transfections gave biologically active material. These data suggest that expression levels of BMP-3 are substantially lower than those of BMP-2A and BMP-2B using native precursors.

Next, the role of the precursor region on formation and secretion of mature BMP-2B
15 was examined. An expression plasmid containing DNA encoding the N-terminal prodomain of BMP-2A spliced to the C-terminal-mature-growth factor domain of BMP-2B (the sequence of which is shown in Figure 2) was assembled. This hybrid BMP-2A/2B construct codes for a protein of 400 amino acids, consisting of residues 1-268 from BMP-2A and residues 277-409 of BMP-2B. The hybrid was assembled from the BMP-2A plasmid (pRK5.bmp2a) by
20 removing the region from the BalI site to the HindIII site and replacing it with the corresponding BalI to HindIII fragment from the BMP-2B plasmid (pRK5.bmp2b).

The resulting expression plasmid (designated pRK5.bmp2/4-1.1) is shown in Figure 3A. Nucleotide sequencing revealed two differences in the BMP-2A sequence compared to that reported by Wozney et al., *supra*: a substitution of A for G at base 261 relative to the ATG
25 start codon, which is silent, and an A for T substitution at base 570 that results in an Arg instead of a Thr at residue 190. (The sequences in Figures 1 and 2 do not reflect the newly found difference at position 190.) The 2A/2B insert sequence is shown in Figure 3B. *E. coli* MM294 cells transformed with this plasmid (*E. coli* MM294/pRK5.bmp2/4-1.1) were deposited with the American Type Culture Collection on May 23, 1990 under ATCC
30 Accession No. 68,330.

pRK5.bmp2/4-1.1, as well as pRK5.bmp2a and pRK5.bmp2b for comparative purposes, were used to transfect 293 cells using the same procedure as described above, and the transfected cells were metabolically labeled using the same procedure as described above, except that they were labeled for four hours with 250 μ Ci/ml each of the 35 S-labeled
35 methionine and cysteine. They were then applied to a 10% SDS-PAGE gel (reduced) using the procedure described above. Figure 4 is the fluorogram exposed for 12 hours at -80°C of this gel (reduced) of conditioned media (5 μ l/lane) from the 293 cells transfected with

plasmids containing either BMP-2A (lane 1), BMP-2A/2B (lane 2), BMP-2B (lane 3), control pRK5 plasmid (lane 4), or no plasmid (lane 5).

For the hybrid, strong bands were found at 36 kD and 23 kD corresponding to the pro and mature forms, respectively. The full-length BMP-2A construct expressed mostly the 36-
5 kD band of the pro form with a small amount of the 18-kD mature form, while for the full-length BMP-2B construct, only a small amount of the 23-kD mature band was found. Thus, greatly enhanced expression of the DNA encoding the BMP-2B mature dimer was observed over expression with the native prod domains.

Biologically active recombinant BMP-2B homodimers were purified from 3-10 liters of
10 conditioned medium from 293 cultures (in 150-mm dishes) transiently transfected with pRK5.bmp2/4-1.1 and DNA encoding the VA RNA gene (Thimmappaya et al., *supra*) as described above but using 28 μ g pRK5.bmp2/4-1.1 and 8 μ g VA gene per dish. One hour after glycerol shock, the media was replaced with serum-free medium [F12:DMEM (1:1) supplemented with 5 μ g/ml human transferrin, 10 μ g/ml insulin, and optionally 10 ng/ml epidermal growth factor, Mather, Biol. Reprod., 23: 243 (1980)] (20 ml of media in each
15 plate). The cells were incubated for 24 hours; the media was harvested, and then fresh medium was added; the cells were incubated again for 24 hours, the medium was harvested and fresh medium was added; and this cycle was repeated once again for a total of three harvests at 24, 48, and 72 hours.

20 Under the conditions of harvesting, the BMP-2B accumulates in the medium to about 200 ng/ml, while background protein levels remain relatively low, as estimated by the intensity of silver-stained SDS-PAGE gels of the conditioned medium. The protein was purified as follows: A 30-ml heparin-Sepharose CL6B column (Pharmacia) was initially equilibrated with 4 M urea, 20 mM TrisCl at pH 7. Then the conditioned medium in 4 M urea,
25 20 mM TrisCl, pH 7, was loaded on the column. The fractions were eluted with a 500-ml gradient of 0 to 0.5 M NaCl in 4 M urea, 20 mM TrisCl, pH 7. One major protein band appeared on the SDS-PAGE gel of the pooled fractions, with an estimated 70-80% purity.

The pooled fractions were concentrated with an Amicon Centricon[®] 10 concentrator about 10-fold, then diluted about 10-fold with 4 M urea, 20 mM Tris, pH 7. The diluted
30 material was loaded onto a 1-ml Pharmacia Mono-Q HR 5/5 column and was eluted with a 0 to 0.3 M NaCl gradient (30 ml) in 4 M urea, 20 mM Tris, pH 7. The peak fractions were pooled, and determined to be about 95% pure by SDS-PAGE. The pooled fractions were dialyzed against 0.1 M acetic acid, lyophilized, and redissolved in 1 ml of 0.1 M acetic acid.

In cases where the purity of the Mono-Q column eluate was judged unsatisfactory, an
35 additional HPLC purification step was employed. This step involved loading the pooled fractions from the Mono-Q column directly on a Vydac C4 RP-HPLC column (100 x 2.1 mm). The HPLC column was eluted with a 30-ml gradient of 0 to 40% N-propanol, 0.1 to 0.06% trifluoroacetic acid. The pooled material from this third step was approximately 95% pure,

as judged by SDS gel electrophoresis. This material was lyophilized and redissolved as described above. Final yield of purified mature BMP-2B was determined by quantitative amino acid analysis; the preparation with the three steps yielded 10 μg /liter of conditioned medium, or approximately 5% overall based on SDS gel analysis.

5 N-terminal amino acid sequencing of the purified mature BMP-2B showed a single amino terminal sequence beginning at residue 285 of BMP-2A/2B (residue 294 of BMP-2B). Sequence data was collected for 18 cycles, and matched exactly that shown underlined in Fig. 3B. No minor sequence was observed. The prominent 36-Kd band observed in the SDS gel of the transfected supernatants was identified as the pro region by amino terminal
10 sequencing after transfer to PVDV membranes. Cleavage of the signal sequence between residues 23 and 24 (...LLGGAAG|LVPELGRRKFAAA) was as predicted by the weight matrix method of von Heijne, Nucl. Acid Res., 144: 683-690 (1986). No cleavage at the nearby RRK sequence was observed.

Recombinant BMP-2B is a disulfide-linked dimer, as shown by a decrease in apparent
15 molecular weight on SDS gel electrophoresis from 33 Kd in the absence of reductants to 23 Kd in the presence of reductants. BMP-2B has two consensus sites for N-glycosylation.

The HPLC-purified recombinant BMP-2B was tested in the bone formation assay of Reddi and Sampath, *supra*, along with TGF- β and a control. In this assay the implants placed into rats were 25 mg demineralized bone powder (DBP) or 25 mg guanidine-HCl-extracted
20 DBP reconstituted with 0, 0.5, 2.0, or 6.0 μg of the purified recombinant BMP-2B or 1 μg recombinant mature human TGF- β 1 (U.S. Pat. No. 4,886,747 issued December 12, 1989). The implants were harvested at 12 days, and the calcium content (Fig. 5A) was measured by atomic absorption spectrophotometry and the alkaline phosphatase content (Fig. 5B) was measured by hydrolysis of p-nitrophenyl phosphate. Duplicate experiments of the 0.5 and
25 2.0 doses of BMP-2B indicated by solid and cross-hatched bars in Fig. 5 were performed.

A significant increase in calcium content (even over DBP, which contains some BMP) was seen with the 2 μg dose of BMP-2B, while the 0.5 μg dose was sufficient to increase alkaline phosphatase. After a 12-day harvest, implants of guanidine-HCl-extracted DBP alone or reconstituted with 1 μg of purified recombinant BMP-2B were fixed and mounted without
30 decalcification. Three-micron sections were cut and stained with haematoxylin and eosin. Microscopic examination of these stained sections showed abundant bone formation in implants reconstituted with BMP-2B as indicated by the presence of calcium deposits. Implants reconstituted with vehicle alone did not form bone.

A construct of the BMP-2A prodomain with the BMP-3 mature region prepared as
35 described above (by replacing the small Ball to HindIII fragment of pRK5.bmp2a with the corresponding Ball-HindIII fragment from pRK5.bmp3) was transfected into 293 cells as described above. In this case, the expression level was no better than the expression levels of the native prosequences for BMP-2A and BMP-3. This experiment shows that the BMP-2A

prodomain does not improve expression levels of every member of the entire BMP family, but rather is effective in enhancing expression of DNA encoding the BMP-2 family.

The ability of the heterologous precursor region to improve secretion of the biologically active dimer may reflect a preference of the BMP-2A precursor region for the BMP-2B mature growth factor sequence. It certainly indicates the importance of the precursor region in proper expression and folding of the biologically active mature dimer form in the BMP-2 family.

Deposit of Materials

The following culture has been deposited with the American Type Culture Collection, 12301 Parklawn Drive, Rockville, MD, USA (ATCC):

<u>Strain</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
MM294/pRK5.bmp2/4-1.1	68,330	May 23, 1990

This deposit was made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture for 30 years from the date of deposit. The organism will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures permanent and unrestricted availability of the progeny of the culture to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent application, whichever comes first, and assures availability of the progeny to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 USC §122 and the Commissioner's rules pursuant thereto (including 37 CFR §1.14 with particular reference to 886 OG 638).

In respect of those designations in which a European patent is sought, a sample of the deposited microorganism will be made available until the publication of the mention of the grant of the European patent or until the date on which the application has been refused or withdrawn or is deemed to be withdrawn, only by the issue of such a sample to an expert nominated by the person requesting the sample. (Rule 28(4) EPC)

The assignee of the present application has agreed that if the culture on deposit should die or be lost or destroyed when cultivated under suitable conditions, it will be promptly replaced on notification with a viable specimen of the same culture. Availability of the deposited strain is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the construct deposited, since the deposited embodiment is intended as a single illustration of one aspect of the invention and any constructs that are functionally equivalent are within the

scope of this invention. The deposit of material herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any aspect of the invention, including the best mode thereof, nor is it to be construed as limiting the scope of the claims to the specific illustration that it represents. Indeed, various
5 modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- 5 (i) APPLICANT: GENENTECH, INC.
- (ii) TITLE OF INVENTION: Mammalian Expression of the BMP-2 Family
- 10 (iii) NUMBER OF SEQUENCES: 12
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: Genentech, Inc.
- (B) STREET: 460 Point San Bruno Blvd
- 15 (C) CITY: South San Francisco
- (D) STATE: California
- (E) COUNTRY: USA
- (F) ZIP: 94080
- 20 (v) COMPUTER READABLE FORM:
- (A) MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: patin (Genentech)
- 25 (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER:
- (B) FILING DATE:
- (C) CLASSIFICATION:
- 30 (vii) PRIOR APPLICATION DATA:
- (A) APPLICATION NUMBER: U.S. Ser. No. 07/528,300
- (B) FILING DATE: 24 May 1990
- (viii) ATTORNEY/AGENT INFORMATION:
- 35 (A) NAME: Hasak, Janet E.
- (B) REGISTRATION NUMBER: 28,616
- (C) REFERENCE/DOCKET NUMBER: 623
- (ix) TELECOMMUNICATION INFORMATION:
- 40 (A) TELEPHONE: 415/266-1896
- (B) TELEFAX: 415/952-9881
- (C) TELEX: 910/371-7168

(2) INFORMATION FOR SEQ ID NO:1:

- 45 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 68 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- 50 (D) TOPOLOGY: linear
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
- 55 CGACCATGGT GGCCGGGACC CGCTGTCTTC TAGCGTTGCT GCTTCCCCAG 50
- GTCCTCCTGG GCGGCGCG 68
- 60

(2) INFORMATION FOR SEQ ID NO:2:

- 65 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 60 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

-20-

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

5 AATGAAAAGG TTGTATTAAA GAACTATCAG GACATGGTTG TGGAGGGTTG 50
TGGGTGTCCG 60

10

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 78 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

20

ATGATTCCTG GTAACCGAAT GCTGATGGTC GTTTTATTAT GCCAAGTCCT 50

25

GCTAGGAGGC GCGAGCCATG CTAGTTG 78

30

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

35 (A) LENGTH: 63 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

40

CAGGAGATGG TAGTAGAGGG ATGTGGGTGC CGCTGAGATC AGGCAGTCCT 50

45

TGAGGATAGA CAG 63

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

50 (A) LENGTH: 50 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AGTGTCCCGC AGCGACGCCG GGAGCCGACG CGCCGCGCGG GTACCTAGCC 50

60

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

65 (A) LENGTH: 50 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

5 TACCCTAACA TGACAGTAGA GTCTTGGCCT TGCAGATAAC CTGGCAAAGA 50

(2) INFORMATION FOR SEQ ID NO:7:

10 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Leu Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg Arg
 1 5 10 15
 Lys Phe Ala Ala Ala
 20 20

(2) INFORMATION FOR SEQ ID NO:8:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 396 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Val Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Leu Pro Gln
 1 5 10 15
 Val Leu Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg
 20 25 30
 Arg Lys Phe Ala Ala Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro
 35 40 45
 Ser Asp Glu Val Leu Ser Glu Phe Glu Leu Arg Leu Leu Ser Met
 50 55 60
 Phe Gly Leu Lys Gln Arg Pro Thr Pro Ser Arg Asp Ala Val Val
 65 70 75
 Pro Pro Tyr Met Leu Asp Leu Tyr Arg Arg His Ser Gly Gln Pro
 80 85 90
 Gly Ser Pro Ala Pro Asp His Arg Leu Glu Arg Ala Ala Ser Arg
 95 100 105
 Ala Asn Thr Val Arg Ser Phe His His Glu Glu Ser Leu Glu Glu
 110 115 120
 Leu Pro Glu Thr Ser Gly Lys Thr Thr Arg Arg Phe Phe Phe Asn
 125 130 135
 Leu Ser Ser Ile Pro Thr Glu Glu Phe Ile Thr Ser Ala Glu Leu
 140 145 150
 Gln Val Phe Arg Glu Gln Met Gln Asp Ala Leu Gly Asn Asn Ser
 155 160 165
 Ser Phe His His Arg Ile Asn Ile Tyr Glu Ile Ile Lys Pro Ala
 170 175 180
 Thr Ala Asn Ser Lys Phe Pro Val Thr Ser Leu Leu Asp Thr Arg

		185		190		195									
	Leu	Val	Asn	Gln	Asn	Ala	Ser	Arg	Trp	Glu	Ser	Phe	Asp	Val	Thr
					200					205				210	
5	Pro	Ala	Val	Met	Arg	Trp	Thr	Ala	Gln	Gly	His	Ala	Asn	His	Gly
					215					220				225	
10	Phe	Val	Val	Glu	Val	Ala	His	Leu	Glu	Glu	Lys	Gln	Gly	Val	Ser
					230					235				240	
	Lys	Arg	His	Val	Arg	Ile	Ser	Arg	Ser	Leu	His	Gln	Asp	Glu	His
					245					250				255	
15	Ser	Trp	Ser	Gln	Ile	Arg	Pro	Leu	Leu	Val	Thr	Phe	Gly	His	Asp
					260					265				270	
	Gly	Lys	Gly	His	Pro	Leu	His	Lys	Arg	Glu	Lys	Arg	Gln	Ala	Lys
20					275					280				285	
	His	Lys	Gln	Arg	Lys	Arg	Leu	Lys	Ser	Ser	Cys	Lys	Arg	His	Pro
					290					295				300	
25	Leu	Tyr	Val	Asp	Phe	Ser	Asp	Val	Gly	Trp	Asn	Asp	Trp	Ile	Val
					305					310				315	
	Ala	Pro	Pro	Gly	Tyr	His	Ala	Phe	Tyr	Cys	His	Gly	Glu	Cys	Pro
					320					325				330	
30	Phe	Pro	Leu	Ala	Asp	His	Leu	Asn	Ser	Thr	Asn	His	Ala	Ile	Val
					335					340				345	
	Gln	Thr	Leu	Val	Asn	Ser	Val	Asn	Ser	Lys	Ile	Pro	Lys	Ala	Cys
					350					355				360	
35	Cys	Val	Pro	Thr	Glu	Leu	Ser	Ala	Ile	Ser	Met	Leu	Tyr	Leu	Asp
					365					370				375	
40	Glu	Asn	Glu	Lys	Val	Val	Leu	Lys	Asn	Tyr	Gln	Asp	Met	Val	Val
					380					385				390	
	Glu	Gly	Cys	Gly	Cys	Arg									
					395	396									

45 (2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 409 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

50 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

55	Met	Ile	Pro	Gly	Asn	Arg	Met	Lys	Met	Val	Val	Leu	Leu	Cys	Gln
	1				5					10				15	
	Val	Leu	Leu	Gly	Gly	Ala	Ser	His	Ala	Ser	Leu	Ile	Pro	Glu	Thr
					20					25				30	
60	Gly	Lys	Lys	Lys	Val	Ala	Glu	Ile	Gln	Gly	His	Ala	Gly	Gly	Arg
					35					40				45	
	Arg	Ser	Gly	Gln	Ser	His	Glu	Lys	Lys	Arg	Asp	Phe	Glu	Ala	Thr
65					50					55				60	
	Leu	Leu	Gln	Met	Phe	Gly	Leu	Arg	Arg	Arg	Pro	Gln	Pro	Ser	Lys
					65					70				75	

-24-

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 400 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

10 Met Val Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Leu Pro Gln
 1 5 10 15

Val Leu Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg
 20 25 30

15 Arg Lys Phe Ala Ala Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro
 35 40 45

20 Ser Asp Glu Val Leu Ser Glu Phe Glu Leu Arg Leu Leu Ser Met
 50 55 60

Phe Gly Leu Lys Gln Arg Pro Thr Pro Ser Arg Asp Ala Val Val
 65 70 75

25 Pro Pro Tyr Met Leu Asp Leu Tyr Arg Arg His Ser Gly Gln Pro
 80 85 90

30 Gly Ser Pro Ala Pro Asp His Arg Leu Glu Arg Ala Ala Ser Arg
 95 100 105

Ala Asn Thr Val Arg Ser Phe His His Glu Glu Ser Leu Glu Glu
 110 115 120

35 Leu Pro Glu Thr Ser Gly Lys Thr Thr Arg Arg Phe Phe Phe Asn
 125 130 135

Leu Ser Ser Ile Pro Thr Glu Glu Phe Ile Thr Ser Ala Glu Leu
 140 145 150

40 Gln Val Phe Arg Glu Gln Met Gln Asp Ala Leu Gly Asn Asn Ser
 155 160 165

Ser Phe His His Arg Ile Asn Ile Tyr Glu Ile Ile Lys Pro Ala
 170 175 180

45 Thr Ala Asn Ser Lys Phe Pro Val Thr Arg Leu Leu Asp Thr Arg
 185 190 195

50 Leu Val Asn Gln Asn Ala Ser Arg Trp Glu Ser Phe Asp Val Thr
 200 205 210

Pro Ala Val Met Arg Trp Thr Ala Gln Gly His Ala Asn His Gly
 215 220 225

55 Phe Val Val Glu Val Ala His Leu Glu Glu Lys Gln Gly Val Ser
 230 235 240

Lys Arg His Val Arg Ile Ser Arg Ser Leu His Gln Asp Glu His
 245 250 255

60 Ser Trp Ser Gln Ile Arg Pro Leu Leu Val Thr Phe Gly His Asp
 260 265 270

Gly Arg Gly His Ala Leu Thr Arg Arg Arg Arg Ala Lys Arg Ser
 275 280 285

65 Pro Lys His His Ser Gln Arg Ala Arg Lys Lys Asn Lys Asn Cys
 290 295 300

Arg Arg His Ser Leu Tyr Val Asp Phe Ser Asp Val Gly Trp Asn
 305 310 315

5 Asp Trp Ile Val Ala Pro Pro Gly Tyr Gln Ala Phe Tyr Cys His
 320 325 330

Gly Asp Cys Pro Phe Pro Leu Ala Asp His Leu Asn Ser Thr Asn
 335 340 345

10 His Ala Ile Val Gln Thr Leu Val Asn Ser Val Asn Ser Ser Ile
 350 355 360

Pro Lys Ala Cys Cys Val Pro Thr Glu Leu Ser Ala Ile Ser Met
 365 370 375

15 Leu Tyr Leu Asp Glu Tyr Asp Lys Val Val Leu Lys Asn Tyr Gln
 380 385 390

20 Glu Met Val Val Glu Gly Cys Gly Cys Arg
 395 400

(2) INFORMATION FOR SEQ ID NO:11:

25 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 56 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

30 Lys Arg His Val Arg Ile Ser Arg Ser Leu His Gln Asp Glu His
 1 5 10 15

35 Ser Trp Ser Gln Ile Arg Pro Leu Leu Val Thr Phe Gly His Asp
 20 25 30

Gly Lys Gly His Pro Leu His Lys Arg Glu Lys Arg Gln Ala Lys
 35 40 45

40 His Lys Gln Arg Lys Arg Leu Lys Ser Ser Cys
 50 55 56

(2) INFORMATION FOR SEQ ID NO:12:

45 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 60 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Gly Gln His Val Arg Ile Ser Arg Ser Leu Pro Gln Gly Ser Gly
 1 5 10 15

55 Asn Asn Ala Gln Leu Arg Pro Leu Leu Val Thr Phe Gly His Asp
 20 25 30

Gly Arg Gly His Ala Leu Thr Arg Arg Arg Arg Ala Lys Arg Ser
 35 40 45

60 Pro Lys His His Ser Gln Arg Ala Arg Lys Lys Asn Lys Asn Cys
 50 55 60

WHAT IS CLAIMED IS:

1. A DNA construct comprising DNA encoding a mature BMP-2 upstream of which is DNA encoding a precursor portion of a mammalian protein other than that of BMP-2.
2. The DNA construct of claim 1 wherein the BMP-2 is human BMP-2.
3. The DNA construct of claim 2 wherein the BMP-2 is human BMP-2B.
4. The DNA construct of claim 1 wherein the precursor portion comprises a signal sequence.
5. The DNA construct of claim 1 wherein the precursor portion has at least 25% amino acid sequence identity to the native precursor portion of the BMP-2 from the N-terminus of the BMP-2 precursor to the first cysteine residue in the mature region of the BMP-2.
6. The DNA construct of claim 5 wherein the precursor portion is from a different BMP-2.
7. The DNA construct of claim 6 wherein the mature BMP-2 is mature BMP-2B and the precursor portion is from BMP-2A.
8. The DNA construct of claim 7 wherein the BMP-2A and BMP-2B are human BMPs.
9. ~~An expression vector comprising the DNA construct of claim 1.~~
10. ~~An expression vector comprising the DNA construct of claim 5.~~
11. ~~An expression vector comprising the DNA construct of claim 7.~~
12. ~~The expression vector of claim 11 that is pRK5.bmp2/4-1.1.~~
13. ~~A mammalian host transformed with the expression vector of claim 9.~~
14. ~~A mammalian host transformed with the expression vector of claim 10.~~
15. ~~A mammalian host transformed with the expression vector of claim 11.~~
16. ~~A mammalian host transformed with the expression vector of claim 12.~~
17. ~~An *E. coli* host transformed with the expression vector of claim 12 deposited as ATCC No. 68,330.~~
18. ~~The host of claim 13 that is a 293 or Chinese hamster ovary cell line.~~
19. In a method for expressing DNA encoding a BMP-2 in mammalian cells, the improvement which comprises employing as the host the host of claim 13.
20. In a method for expressing DNA encoding a BMP-2 in mammalian cells, the improvement which comprises employing as the host the host of claim 14.
21. In a method for expressing DNA encoding a BMP-2 in mammalian cells, the improvement which comprises employing as the host the host of claim 15.
22. In a method for expressing DNA encoding mature BMP-2B in mammalian cells, the improvement which comprises employing as the host the host of claim 16.

```

10      20      30      40      50      60      70      80      90
bmp2a  *MVAGTRCLLALLPQVLLGGA--AGLVPELGRRKFA--AASSGRPSSQPSDEVLSEFELRLLSMFGLKQRTPSRD VVPPYMLDLYRRHSGQP GSPAPD
10      20      30      40      50      60      70      80      90
bmp2b  *MIPGNRMKMWVLLCQVLLGGASHASLIPETGKKKVAEIQGHAGRRRGGQSHEKRDFFEA TLLQMFGLRRRPPQPSKSAVIPDYMRDLYRLQSGEEEEEQIH
100     110     120     130     140     150     160     170     180     190
bmp2a  *HR-L---ERAASRANTVRSFHHHEESLEELPETSCKTTRRRFFNLSSIPTEEFITSAELQVFRQMDALGNSSFFHHRINIYEI IK-PATANSKFPV TSL
110     120     130     140     150     160     170     180     190
bmp2b  *STGLEYPERPASRANTVRSFHHHEEHL ENIPGTSEN S AFRFLNLS S I PENEVIS S AELRLFR EQVDQGP DW ERGF--HRINIYEV MKPPAEVVP GHL ITRL
200     210     220     230     240     250     260     270     280     290
bmp2a  *LDTRLVNQNASRWESFDVTPAVMRWTAGGHANHGFFVEVAHLEEKQGVSKRHVRI SRS L HQDEHSWSQIRPLLVTF GHDGKGHPL--HKREKRQAK-HKQ
200     210     220     230     240     250     260     270     280     290
bmp2b  *LDTRLVHNVTRETFDVS PAVLRWTR EKQPNYGLAIEVTHLHQTRTHQGQHVRI SRS L POGSGNNAQLRPLLVTF GHDGRGHALTRRRRAKRS PKHHSQ
290     300     310     320     330     340     350     360     370     380
bmp2a  *RKRLKS-SCKRHPLYVDESDVGWNDWI VAPPGYHAFYCHGECPPFLADHLNSTNHAI VQTLYNSVNSKIPKACCVPTELSAI SMLYLDENEKVV LKNYQD
290     300     310     320     330     340     350     360     370     380
bmp2b  *RARKKNKNCRRHSLYVDESDVGWNDWI VAPPGYQAFYCHGD C PFLADHLNSTNHAI VQTLYNSVNS S I PKACCVPTQLSAI SMLYLDEYDKVVLKNYQD
390
bmp2a  *MVVEGCGCR
400     410     420     430     440     450     460     470     480     490
bmp2b  *MVVEGCGCR

```

FIG. 1

FIG. 2

BMP 2A/2B

10	20	30	40	50	60	70
MVAGTRCLLALLLPQVLLGGAAGLVPELGRRKFAAASSGRPSSQPSDEVLSEFELRLLSMFGLKQRTPS						
80	90	100	110	120	130	140
RDAVPPYMLDLYRRHSGQPGSPAPDHLERAASRANTVRSFHHHEESLEELPETSCKTTRRRFFNLSSIP						
150	160	170	180	190	200	210
TEEFITSAELQVREQMODALGNSSFHHRINIYEI IKPATANSKFPVTRLLDTRLVQNQASRWESFDVT						
220	230	240	250	260	270	280
PAMRWTAQGHANHGFFVEVAHLEEKQGVSKRRHVRISRSLHQDEHSWSQIRPLLVTFFGHDRGHALTRRR						
290	300	310	320	330	340	350
RAKRSPKHHSQARAKKKNKNCRRHSLSLYVDFSDVGWNDWIVAPPGYQAFYCHGDCFPPLADHLNSTNHAIVQ						
360	370	380	390	400		
TLVNSVNSSIPKACCVPTELSAISMLYLDEYDKVVLKKNYQEMVVEGGCR						